homework7

xiao-qiyuan

October 24, 2018

Read data

```
marvel = read.csv('F:\\UIUC\\590DV\\homework7\\marvel-wikia-data.csv',header
= TRUE, stringsAsFactors = FALSE)
dc = read.csv('F:\\UIUC\\590DV\\homework7\\dc-wikia-data.csv',header = TRUE,
stringsAsFactors = FALSE)
```
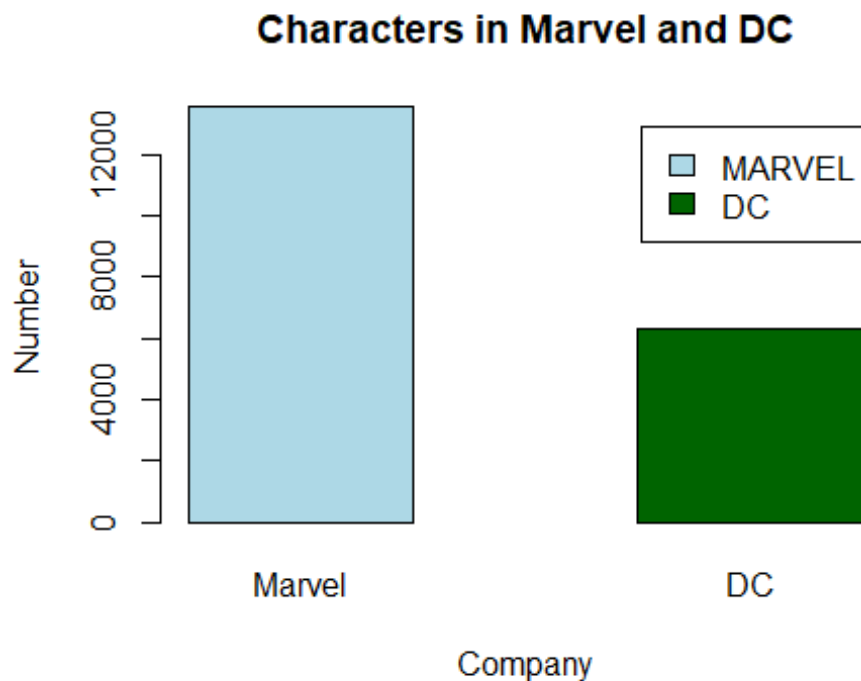
Who wins?

Absolutely MARVEL!

As for who wins among Marvel and DC, I believe we need to know the percentage of how many bad guys and good guys in those universe so that we can how many bad guy on average do a good guy fight against.

Because the rows in two data frame are different as well as the type of align, we need to create a new data frame to analysis.

```
mg = length(marvel$ALIGN[which(marvel$ALIGN=='Good Characters')])
mb = length(marvel$ALIGN[which(marvel$ALIGN=='Bad Characters')])
mn = length(marvel$ALIGN[which(marvel$ALIGN=='Neutral Characters')])
dg = length(dc$ALIGN[which(dc$ALIGN=='Good Characters')])
db = length(dc$ALIGN[which(dc$ALIGN=='Bad Characters')])
dn = length(dc$ALIGN[which(dc$ALIGN=='Neutral Characters')])
dt = length(dc$ALIGN[which(dc$ALIGN=='Reformed Criminals')])
md = data.frame(m = c(mg,mb,mn,mt = c(0)),d = c(dg,db,dn,dt),stringsAsFactors
= FALSE)
```

Now, let's take a look at how many good and bad characters are in Marvel and DC. First, Let's take a look at how many characters are in those company

```
a = c(sum(md$m),sum(md$d))
count1= matrix(a,nrow = 1,ncol = 2 )
barplot(count1,
        col = c('lightblue','darkgreen'),
        main = 'Characters in Marvel and DC',
        xlab = 'Company',ylab = 'Number',beside = TRUE,names.arg =
c('Marvel','DC'),legend.text = c('MARVEL','DC'))
```
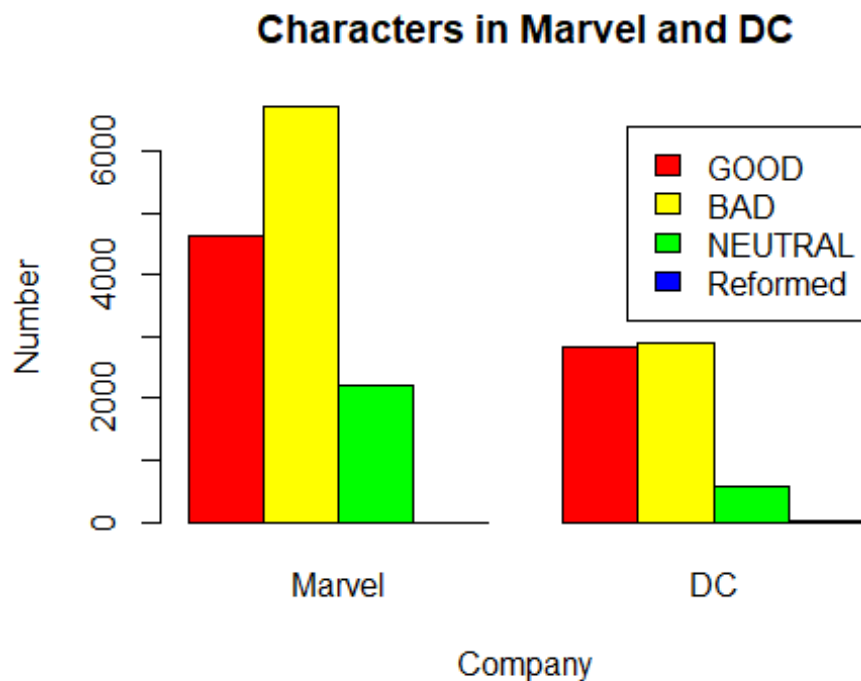
Characters in Marvel and DC

We can see that Marvel has more character than DC, indicating Marvel is better.

Then, lets see how many of each good, bad and neutral character.

```
b = c(md$m,md$d)
counts = matrix(b,nrow = 4,ncol = 2 )
barplot(counts,
        col = c('red','yellow','green','blue'),
        main = 'Characters in Marvel and DC',
        xlab = 'Company',ylab = 'Number',beside = TRUE,names.arg =
c('Marvel','DC'),legend.text = c('GOOD','BAD','NEUTRAL','Reformed'),width =
0.01)
```
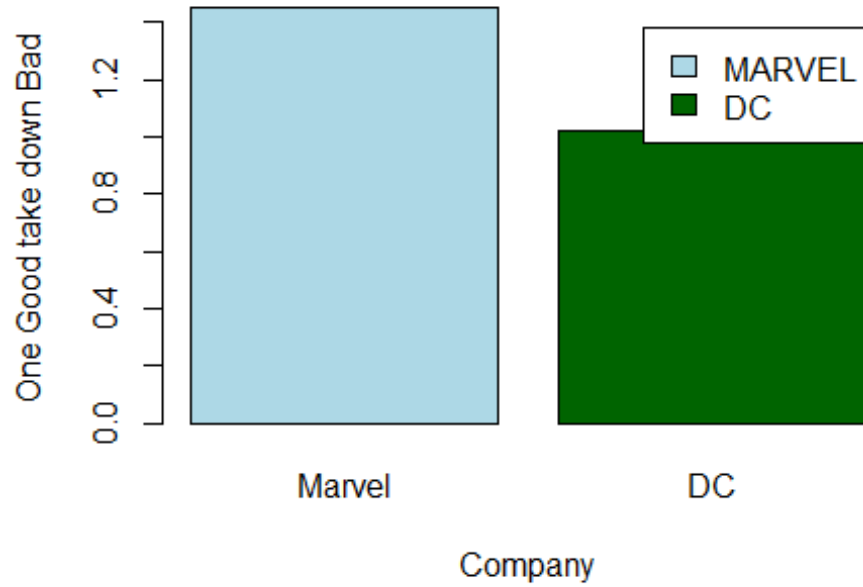
Characters in Marvel and DC

We see Marvel has more bad characters and more good characters than DC

So, one god character need to take down how many bad charcter on average in those two companies?

```r
d = c(mb/mg,db/dg)
barplot(d,
        col = c('lightblue','darkgreen'),
        main = 'Characters in Marvel and DC',
        xlab = 'Company',ylab = 'One Good take down Bad',beside =
TRUE,names.arg = c('Marvel','DC'),legend.text = c('MARVEL','DC'),width = 0.1)
```

Characters in Marvel and DC

Thus, good characters in marvel need to take down more bad guys than in DC. Hence, heros in Marvel are better than heros in DC. Thus, Marvel wins!